# CRITICAL SELF-ORGANIZATION OF ASTROPHYSICAL SHOCKS

M. A. Malkov, P. H. Diamond
University of California at San Diego, 9500 Gilman Dr, La Jolla, CA 92093-0319, USA
*mmalkov@physics.ucsd.edu; pdiamond@physics.ucsd.edu*

AND

H. J. Völk
Max-Planck Institut für Kernphysik, D-69029, Heidelberg, Germany
*Draft version February 1, 2008*

## ABSTRACT

There are two distinct regimes of the first order Fermi acceleration at shocks. The first is a linear (test particle) regime in which most of the shock energy goes into thermal and bulk motion of the plasma. The second is an efficient regime when it goes into accelerated particles. Although the transition region between them is narrow, we identify the factors that drive the system to a *self-organized critical state* between those two. Using an analytic solution, we determine this critical state and calculate the spectra and maximum energy of accelerated particles.

*Subject headings:* acceleration of particles — cosmic rays — diffusion — hydrodynamics — radiation mechanisms:non-thermal — shock waves

## 1. INTRODUCTION

It is now generally recognized that most of the observed gamma radiation derives in one way or another from accelerated particles. Radio and x-ray spectral components from a variety of astrophysical objects are believed to have a similar origin. High energy neutrinos, whose detection is on the program for the future and existing water/ice detectors, must also be related to the ultra-high energy cosmic rays (UHECR). Their origin is, in turn, a mystery and the huge Auger detector complex is now being built to elucidate it (Blandford 1999, Cronin 1999).

There has been essential progress in our understanding of how accelerated particles produce the radiation detected. The models concentrate on the synchrotron and inverse Compton emission for electrons and on the $\gamma$ ray and neutrino production in $pp$ and $p\gamma$ reactions for protons. However, the primary spectrum of accelerated particles remains a stumbling block making predictions of otherwise similar models so different (e.g., Mannheim *et al.*, 1999, Waxman & Bahcall 1999).

The "standard" mechanism of particle acceleration, capable of producing nonthermal power-law spectra extending over many decades in energy is the I-order Fermi or diffusive shock acceleration. It was originally suggested to explain the origin of galactic cosmic rays (CRs). For the purposes of the high-energy radiation and UHECR it is usually adopted as an axiom, and mostly only in its simplest, test particle (TP) or linear realization. In particular, it is assumed that any strong nonrelativistic shock routinely produces a $E^{-2}$ spectrum of protons and/or electrons. In fact this spectrum arises from a simple formula $F \sim E^{-(r+2)/(r-1)}$ where $r$ is the shock compression (for $r = 4$). However, it is valid only if the shock thickness is much smaller than the particle mean free path. This, in turn, is true only if the energy content of accelerated particles is small compared to the shock energy (inefficient acceleration) so that the shock structure is maintained by the thermal, not by the high energy particles. Otherwise,

the accelerated particles create the shock structure on their own and if so, then obviously on a scale that is larger or of the order of their mean free path, thus making the above formula invalid. Therefore, the TP regime requires a very low CR number density (the rate of injection $\nu$ into the acceleration process), which appears to be impossible in the parameter range of interest. It has been inferred from observations (e.g., Lee 1982), simulations (e.g., Bennet & Ellison 1995), and theory (e.g., Malkov 1998) that the CR number density $n_{\rm CR}$ at a strong shock must be $\sim 10^{-3}$ of the background density $n_1$ upstream. It is important to emphasize here that when the actual injection rate $\nu$ exceeds the critical value (denote it $\nu_2$), the test particle ($E^{-2}$) solution simply does not exist. Simple measures, such as calculating corrections, are intrinsically inadequate. What happens is that two other solutions with considerably higher efficiencies branch off at a somewhat lower injection rate $\nu_1 < \nu_2$, one of which disappears again at $\nu = \nu_2$, together with the test particle solution.

Thus, it seems to be difficult to put an accelerating shock into a regime in which the CR energy production rate (acceleration efficiency) could be gradually adjusted by changing parameters. It is either too low (TP regime) or it is close to unity. Note, that this situation is quite suggestive of that occurring in phase transitions or bifurcations.

Generally, neither of those extreme regimes provide an adequate description of particle spectra and related emission. Nevertheless, we argue in this *Letter* that despite this apparent lack of regulation ability, shocks must be still capable of *self-regulation* and *self-organization*. The transition region between the two acceleration regimes (critical region) is very narrow in control parameters like $\nu$. On the other hand, the self-regulation can work efficiently only when the parameters are within this region. This requirement determines them, and also resolves the question of the mechanism of self-regulation. The above consideration is similar to the concept of *self-organized criticality* (SOC)

(e.g., Bak, Tang & Wiesenfeld 1987, Hwa & Kardar 1992, Diamond & Hahm 1995).

## 2. FORMULATION OF THE PROBLEM

We use the diffusion-convection equation (e.g., Drury 1983) for describing the distribution of high energy particles (CRs). We assume that the gaseous discontinuity (also called the subshock) is located at $x = 0$ and the shock propagates in the positive $x$- direction. Thus, the flow velocity in the shock frame can be represented as $V(x) = -u(x)$ where the (positive) flow speed $u(x)$ jumps from $u_2 \equiv u(0-)$ downstream to $u_0 \equiv u(0+) > u_2$ across the subshock and then gradually increases up to $u_1 \equiv u(+\infty) \geq u_0$. In a steady state the equation reads

$$u\frac{\partial f}{\partial x} + \kappa(p)\frac{\partial^2 f}{\partial x^2} = \frac{1}{3}\frac{du}{dx}p\frac{\partial f}{\partial p}, \qquad (1)$$

where $f(x,p)$ is the isotropic (in the local fluid frame) part of the particle distribution. This is assumed to vanish far upstream ($f \to 0$, $x \to \infty$), while the only bounded solution downstream is obviously $f(x,p) = f_0(p) \equiv f(0,p)$. The most plausible assumption about the cosmic ray diffusivity $\kappa(p)$ is that of the Bohm type, i.e., $\kappa(p) = Kp^2/\sqrt{1+p^2}$ (the particle momentum $p$ is normalized to $mc$). In other words $\kappa$ scales as the gyroradius, $\kappa \sim r_g(p)$. The reference diffusivity $K$ depends on the $\delta B/B$ level of the MHD turbulence that scatters the particles in pitch angle. The minimum value for $K$ would be $K \sim mc^3/eB$ if $\delta B \sim B$. Note that this plain parameterization of this important quantity is perhaps the most serious incompleteness of the theory which will be discussed later.

To include the backreaction of accelerated particles on the plasma flow three further equations are needed. First, one simply writes the conservation of the momentum flux in the smooth part of the shock transition ($x > 0$, the so-called CR-precursor)

$$P_c + \rho u^2 = \rho_1 u_1^2, \quad x > 0 \qquad (2)$$

where $P_c$ is the pressure of the high energy particles

$$P_c(x) = \frac{4\pi}{3}mc^2 \int_{p_0}^{p_1} \frac{p^4 dp}{\sqrt{p^2+1}} f(p,x) \qquad (3)$$

It is assumed here that there are no particles with momenta $p > p_1$ (they leave the shock vicinity because there are no MHD waves with sufficiently long wave length $\lambda$, since the cyclotron resonance requires $p \sim \lambda$). The momentum region $0 < p < p_0$ cannot be described by equation (1) and the behavior of $f(p)$ at $p \sim p_0$ is described by the injection parameters $p_0$ and $f(p_0)$ (Malkov 1997, [M97]). The plasma density $\rho(x)$ can be eliminated from equation (2) by using the continuity equation $\rho u = \rho_1 u_1$. Finally, the subshock strength $r_s$ can be expressed through the Mach number $M$ at $x = \infty$ (e.g., Landau & Lifshitz)

$$r_s \equiv \frac{u_0}{u_2} = \frac{\gamma+1}{\gamma-1+2R^{\gamma+1}M^{-2}} \qquad (4)$$

where the precursor compression $R \equiv u_1/u_0$ and $\gamma$ is the adiabatic index of the plasma.

The system of equations (1,2,4) describes in a self-consistent manner the particle spectrum and the flow structure. An efficient way to solve it is to reduce this system to one integral equation (M97). A key dependent variable is an integral transform of the flow profile $u(x)$ with a kernel suggested by an asymptotic solution of the system (1)-(2) which has the form

$$f(x,p) = f_0(p)\exp\left[-\frac{q}{3\kappa}\Psi\right]$$

where

$$\Psi = \int_0^x u(x')dx'$$

and the spectral index downstream $q(p) = -d\ln f_0/d\ln p$. The integral transform is as follows

$$U(p) = \frac{1}{u_1}\int_{0-}^{\infty}\exp\left[-\frac{q(p)}{3\kappa(p)}\Psi\right]du(\Psi) \qquad (5)$$

and it is related to $q(p)$ through the following formula

$$q(p) = \frac{d\ln U}{d\ln p} + \frac{3}{r_s R U(p)} + 3 \qquad (6)$$

Thus, once $U(p)$ is found both the flow profile and the particle distribution can be determined by inverting transform (5) and integrating equation (6). Now, using the linearity of equation (2) ($\rho u = const$), we derive the integral equation for $U$ by applying the transformation (5) to the $x-$ derivative of equation (2) (M97). The result reads

$$U(t) = \frac{r_s-1}{Rr_s} + \frac{\nu}{Kp_0}\int_{t_0}^{t_1}dt'\left[\frac{1}{\kappa(t')} + \frac{q(t')}{\kappa(t)q(t)}\right]^{-1}$$
$$\times \frac{U(t_0)}{U(t')}\exp\left[-\frac{3}{Rr_s}\int_{t_0}^{t'}\frac{dt''}{U(t'')}\right] \qquad (7)$$

where $t = \ln p$, $t_{0,1} = \ln p_{0,1}$. Here the injection parameter

$$\nu = \frac{4\pi}{3}\frac{mc^2}{\rho_1 u_1^2}p_0^4 f_0(p_0) \qquad (8)$$

is related to $R$ by means of the following equation

$$\nu = Kp_0\left(1 - R^{-1}\right)$$
$$\times \left\{\int_{t_0}^{t_1}\kappa(t)dt\frac{U(t_0)}{U(t)}\exp\left[-\frac{3}{Rr_s}\int_{t_0}^{t}\frac{dt'}{U(t')}\right]\right\}^{-1} \qquad (9)$$

The equations (4,7,9) form a closed system that can be easily solved numerically. We analyze the results in the next section.

## 3. MECHANISMS OF CRITICAL SELF-ORGANIZATION

The critical nature of this acceleration process is best seen in variables $R, \nu$. The quantity $R - 1$ is a measure of shock modification produced by CRs, in fact $(R-1)/R = P_c(0)/\rho_1 u_1^2$ (eq.[2]) and may be regarded as an order parameter. The injection rate $\nu$ characterizes the CR density at the shock front and can be tentatively treated as a control parameter. It is convenient to plot the function $\nu(R)$ instead of $R(\nu)$ (using equation [9]), since $R(\nu)$ is not always a single-valued function, Fig. 1.

The injection rate $\nu$ at the subshock should be calculated given $r_s(R)$ (M97) with the self-consistent determination of the flow compression $R$ on the basis of the $R(\nu)$ dependence obtained. However, in view of its critical character, this solution can be physically meaningful only in regimes far from criticality, i.e., when $R \approx 1$ (test particle regime) or $R \gg 1$ (efficient acceleration). But, it is difficult to see how this system could stably evolve remaining in one of these two regimes. Indeed, if $\nu$ is subcritical it will inevitably become supercritical when $p_1$ is sufficiently high. Once it happened, however, the strong subshock reduction (equation [4]) will reduce $\nu$ and drive the system back to the critical regime, Fig. 2.

The maximum momentum $p_1$ is subject to self-regulation as well. Indeed, when $R \gg 1$, the generation and propagation of Alfven waves is characterized by strong inclination of the characteristics of wave transport equation towards larger wavenumbers $k$ on the $k - x$ plane due to wave compression. Thus, considering particles with $p \lesssim p_1$ inside the precursor, one sees that they are in resonance with waves that must have been excited by particles with $p > p_1$ further upstream but, there are no particles with $p > p_1$. Therefore, the required waves can be excited only locally by the same particles with $p \lesssim p_1$ which substantially diminishes the amplitude of waves that are in resonance with particles from the interval $p_1/R < p < p_1$. (The left inequality arises from the resonance condition $kcp \approx eB/mc$ and the frequency conservation along the characteristics $ku(x) = const$). This will worsen the confinement of these particles to the shock front. The quantitative study of this process is the subject of current research. What can be inferred from Fig. 2 now, is that the decrease of $p_1$ straightens and rise the curve $\nu(R)$, so that it returns to the monotonic behaviour. However once the actual injection becomes subcritical (and thus $R \to 1$) then $p_1$ will grow again restoring the two extrema on the curve $\nu(R)$.

The above dilemma is quite typical for dynamical systems that are close to criticality or marginal stability. A natural way to resolve it consists in collapsing the extrema into an inflection point so that a *self-organized critical* (SOC) acceleration regime is established being determined by the conditions $\nu'(R^*) = \nu''(R^*) = 0$. These conditions not only determine unique critical values $R^*$ and $\nu^* \equiv \nu(R^*)$ but also yield the maximum momentum $p_1$ as a function of $M$, which is shown in Fig. 3. A few particle spectra that develop in the SOC states for different Mach numbers are shown in Fig. 4, along with the asymptotic $M = \infty$ non SOC spectrum. The latter can be calculated in a closed form (M97). Note that the hardening of the spectra in about the last decade below the cut-off is entirely due to the abrupt cut-off itself.

## 4. DISCUSSION

The detailed microphysics behind the SOC is extremely complex and must include the self-consistent turbulence evolution and particle acceleration with their strong back-reaction on the shock structure. We have simplified it and argued that the most important dynamical components, the bulk plasma flow and the high-energy particles, must be in a balance that constitutes a certain equipartition of the shock energy between the two. This was done by identifying the factors that prevent either of them from prevailing alone.

The above situation is similar to that in e.g., a simple sandpile paradigm of the SOC. It is impossible (in fact unnecessary) to describe the individual grain dynamics, but it is clear that when the critical macroscopic characteristic of the system (the slope of the sandpile) becomes too steep due to the action of external factors, like tilting of the entire system or addition of sand at the top, the sandpile relaxes bringing the slope to its critical magnitude.

## 5. POSSIBLE FEEDBACK FROM OBSERVATIONS

Perhaps the most significant observational aspect is the particle spectrum. Although the conversion of detected radiation spectra into the primary particle spectra is ambiguous, in some cases it may be compared with the theory. The most striking prediction is that in shocks with *very high maximum particle energy*, the spectra must be harder than $q = 4$ (or 2 in the normalization $f(E)dE$). This is because of a very low injection requirement for efficient acceleration in such shocks (Fig.1). If we (conservatively) set $n_{CR}/n_1 \sim 10^{-3}$, then $\nu \sim 10^{-3}cp_0/mu_1^2$ which may easily exceed $\nu_2$ (local maximum on the $\nu(R)$ curve) already for $p_1 \gtrsim 10^4 - 10^5$, putting the acceleration into a strongly nonlinear regime. This should have important observational consequences.

First, the particle energy is concentrated at the upper cut-off instead of being evenly distributed over the logarithmic energy bands as in the test particle $E^{-2}$ solution. This makes the upper bounds on CR generated neutrino fluxes (see e.g., Bahcall & Waxman 1999 and Mannheim *et al.,* 1999 and references therein) rather ambiguous. Indeed, the UHECR spectrum is normalized to the observed one at $E_{norm} \sim 10^{19}$ eV while being obscured by the galactic background at the energies $E \lesssim 10^{18}$ eV. Thus, CR spectra harder than $E^{-2}$ imply that the upper limit on the neutrino fluxes e.g., derived by Waxman & Bahcall 1999 should be even lower than the $E^{-2}$ spectrum implies for the energies $E_\nu < 5 \cdot 10^{17}$ eV ($E_\nu/E_{CR} \simeq 0.03$ is a typical energy relation). According to the same logic, it should be increased for higher energies. Note that the upper cut-offs in individual shocks contributing to the UHECR must be still much higher than $E_{norm}$ to validate our simplified handling of particle losses. On the other hand if the sources with $E_{max} < E_{norm}$ contribute significantly, the measurements at $E_{norm}$ tell us nothing about their normalizations and the upper bound on neutrino fluxes may be increased up to the level dictated by the CR observations in the lower energy range, as suggested by Mannheim *et al.,* 1999. This scenario is supported by Fig. 3, provided that there are many strong shocks in the ensemble. The observed CR steep power-law spectrum is then essentially a superposition of flatter or even non-power-law spectra from individual sources properly distributed in $E_{max}$.

To summarize our conclusions, the main factor that should determine the particle primary spectra, and thus the neutrino flux, is how the accelerating shocks are distributed in cut-off momenta, which in a SOC state means in Mach numbers. There is no universal spectral form for individual shocks at the current state of the theory (except a not quite representative case of $M \to \infty$). Therefore, one should understand particle losses mechanisms, since they determine the shock structure and thus the spectra,

4

directly and through $E_{\max}$. These mechanisms are insepa-
rable from the dynamics of strong compressible MHD tur-
bulence generated by those same particles. The further
progress in its study will improve our understanding of
the acceleration process and related radiation.

## REFERENCES

Bahcall & Waxman 1999, hep-ph/9807282
Blandford, R. D. 1999 astro-ph/9906026
Bak, P., Tang, C & Wiesenfeld, K. 1987, Phys. Rev. Lett., 59, 381
Bennett, L. and D. C. Ellison 1995, J. Geophys. Res. 100, A3, 3439
Cronin, J. 1999, Rev. of Modern Phys. , 71, S165
Diamond, P. H., & Hahm, T. S. 1995, Phys. Plasmas, 2, 3640
Drury, L. O'C. 1983, Rep. Prog. Phys. 1983, 46, 973
Hwa T. and Kardar, M. 1992 Phys. Rev. A, 45 (10) 7002

Landau, L. D. and Lifshitz, E. M. 1987 Fluid Mechanics (Oxford: Pergamon )
Lee, M. A. 1982, J. Geophys. Res. 87, 5063
Malkov, M. A. ApJ, 1997, 485, 638 (M97)
Malkov, M. A. Phys. Rev. E, 1998, 58, 4911
Mannheim, K., Protheroe, R. J. and Rachen, J. P., 1999, astro-ph/9812398
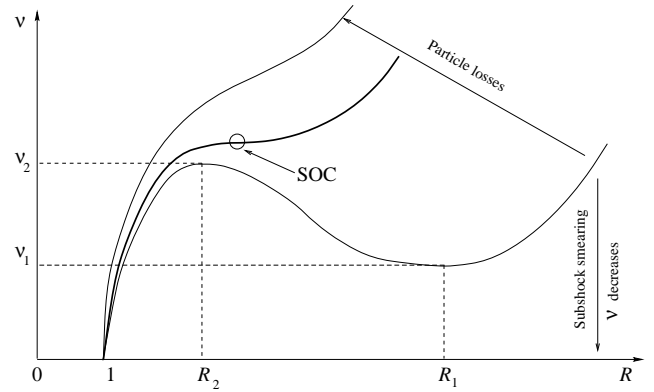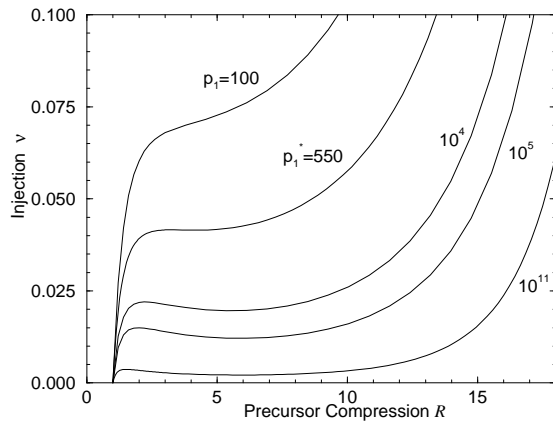Waxman, E., & Bahcall, J., 1999, hep-ph/9807282



FIG. 1.— The nonlinear response of an accelerating shock (characterized by the precursor compression $R$) to the thermal injection $\nu$ given in the form of the function $\nu(R)$ calculated for the fixed injection momentum $p_0 = 10^{-3}$, Mach number $M = 150$ and for different $p_1 = 100; 550; 10^4; 10^5; 10^{11}$. The critical value (see text) $p_1^* = 550$. The TP regime is limited to the region $R \simeq 1$.

FIG. 2.— Bifurcation diagram corresponding to the set of response curves shown in fig. 1. Since $\nu$ and $p_1$ are in reality dynamic rather than control parameters the response curve moves towards the bifurcation curve drawn with the heavy line. It corresponds to the curve marked by $p_1^* = 550$ in Fig. 1.
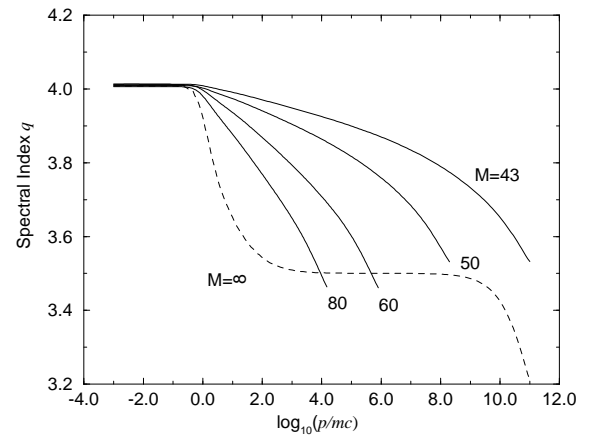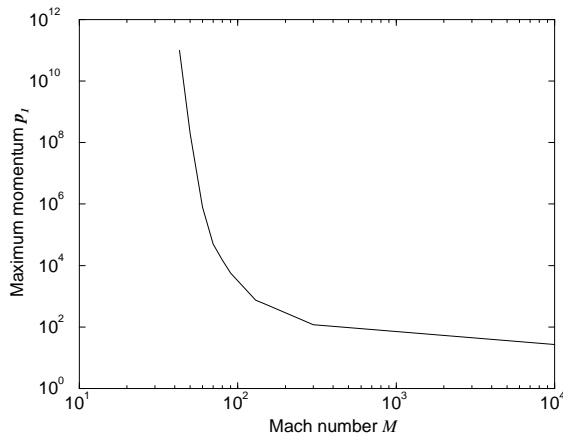


FIG. 3.— Maximum momentum $p_1$ versus Mach number $M$ calculated in the SOC states (see text).

FIG. 4.— The spectral indices in the SOC states shown for different Mach numbers and the corresponding maximum momenta from Fig. 3 (solid curves). For comparison, the dashed curve shows the asymptotic case $M = \infty$, $p_1 = 10^{11}$ and $\nu = 4 \cdot 10^{-5}$. For larger, e.g., SOC injection values this spectrum would have been cut in mildly relativistic region.